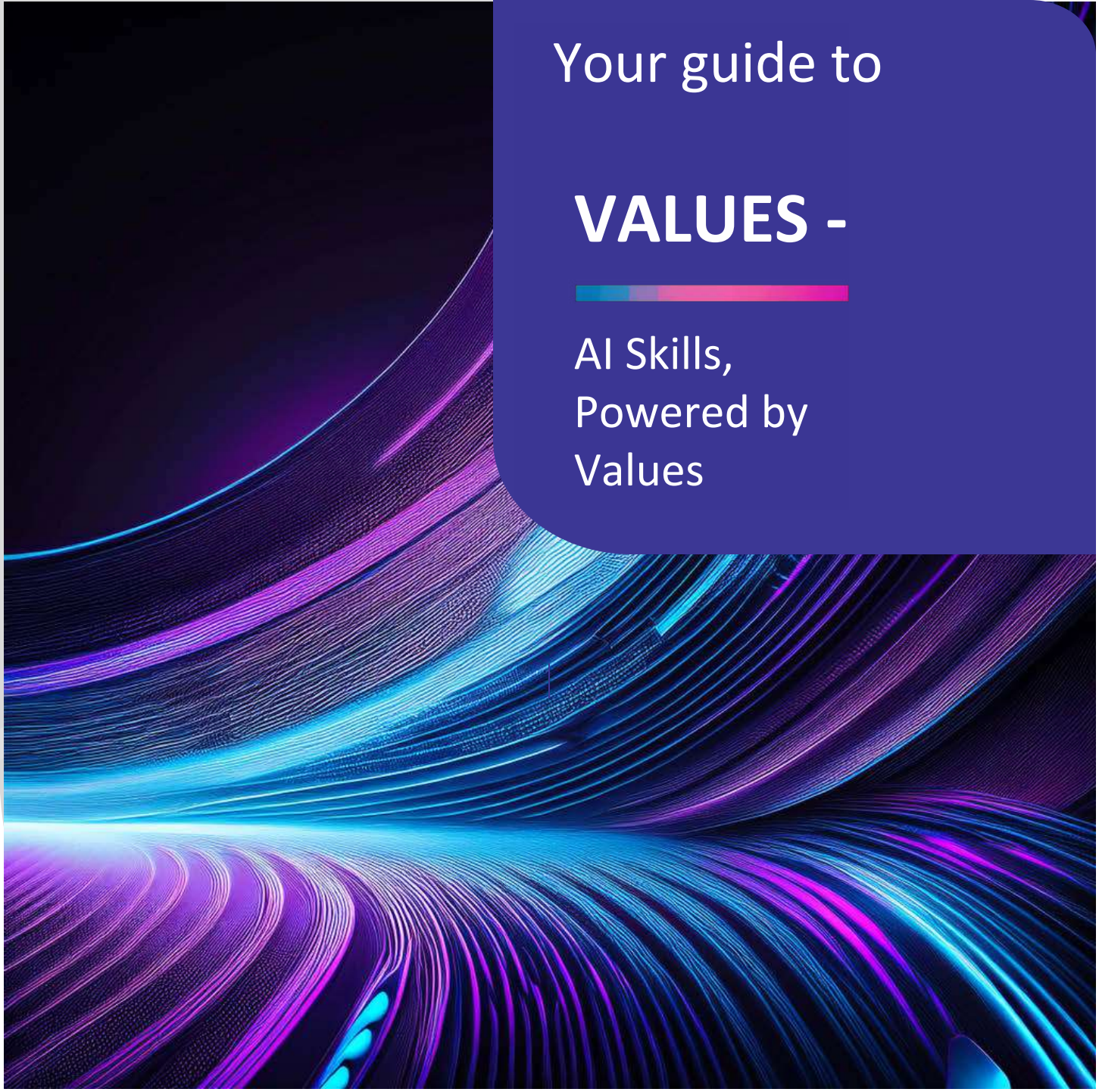


Your guide to

VALUES -

AI Skills, Powered by Values

AI Skills, Powered by Values

values

www.valuesai.eu

2025
Case Studies

By
Generation (Change?)

Co-funded by the European Union

contents

Co-funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Generation (Change?). Neither the European Union nor Generation (Change?) can be held responsible for them.

01

Values: Case Studies guidebook

The Project

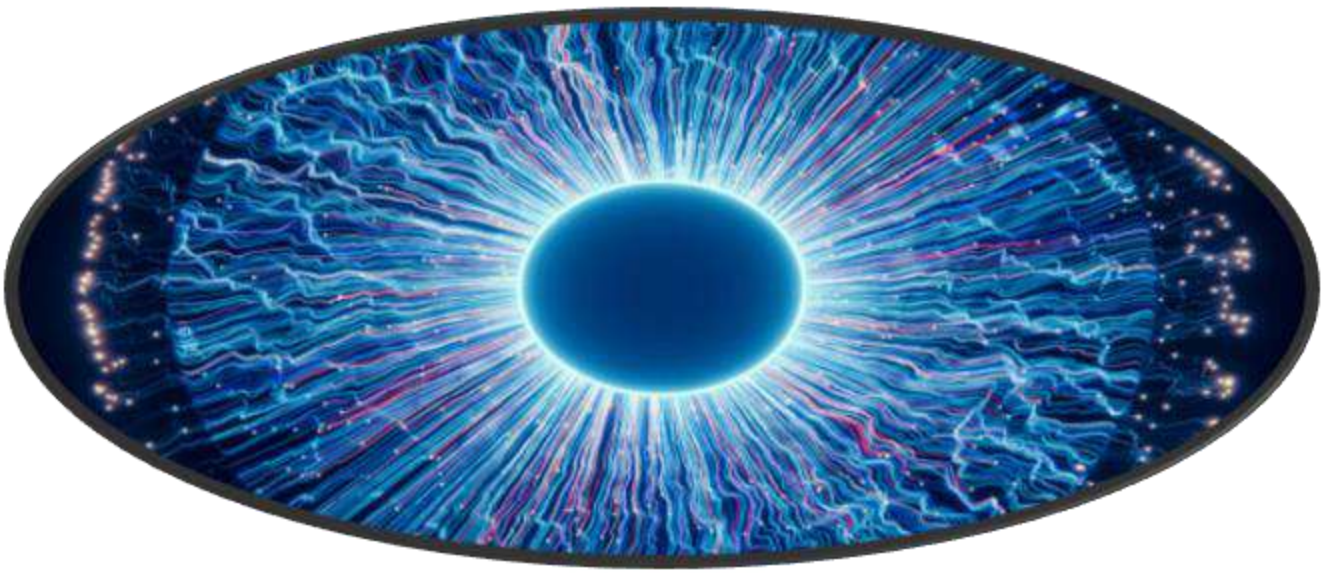
VALUES is an Erasmus+ project dedicated to empowering young Europeans with the knowledge, skills, and ethical understanding needed to thrive in a future shaped by Artificial Intelligence. With AI transforming the way we work, learn, and live, VALUES ensures young people are not just participants but leaders in this digital revolution.

Through cutting-edge training, innovative tools, and a comprehensive **Starter Kit for Ethical AI**, we bridge the gap between technology and humanity. Our interactive **AI Skills Self-Assessment Platform** and dynamic **Open Educational Resources** are designed to make learning accessible, engaging, and impactful.

VALUES isn't just about education—it's about shaping a generation of responsible, values-driven digital citizens who can harness AI ethically to tackle societal challenges, build inclusive communities, and create sustainable solutions.

Project Partners:

- University of Szczecin (US) Lead coordinator
- Stowarzyszenie Wspierania Techniki Polskiej (SWTP)
- Generation (Change?)
- Momentum Marketing Services (MMS)
- European E-Learning Institute (EUEI)
- The Vision Works (tvw GmbH)



Responsible AI is not just about liability - it's about ensuring what you are building is enabling human flourishing

Rumman Chowdhury,
CEO at Parity AI, 2023

Case Study 1

The UK A-levels Algorithm Scandal: A Lesson in Responsible AI in Education

Theme of the Case Study:
- High-stakes exams (A-levels) in the UK
- Social inequality and algorithmic bias
- Responsible AI (RAI) implementation
- Ethics and transparency in educational algorithms

Abstract (Summary): In 2020, due to the cancellation of A-level exams amid the COVID-19 pandemic, the UK government implemented an algorithm to assign student grades. This algorithm, based on historical school performance and ranking data, aimed to standardize results. However, it soon became clear that the system favored students from high-performing, often wealthier schools, at the expense of those from less privileged backgrounds. Many students received grades significantly lower than those predicted by their teachers, sparking nationwide protests and public outrage. Eventually, the government reversed its decision and reinstated teacher-predicted grades.

Key Takeaways:
- Transparency and Auditability: Educational AI systems must be transparent and regularly audited to ensure fairness.
- Context-Aware Design: Algorithms should be developed with attention to social diversity and must avoid reinforcing existing biases.

Relevance to the Reader: This case study highlights the critical importance of designing and deploying AI systems responsibly in education. It illustrates how a lack of transparency, failure to consider social context, and overreliance on historical data can exacerbate existing inequalities. For policymakers, educators, and edtech developers, this serves as a cautionary tale against the uncritical application of algorithms in decision-making processes, especially in sensitive domains like student assessment.

Key Takeaways:
- Stakeholder Involvement: Teachers, students, and communities should be involved in the design and deployment of AI systems.
- Responsiveness and Accountability: Institutions must be ready to respond swiftly to unintended consequences of algorithmic decisions and correct them promptly.

Case Study 2

Biased Amazon Recruitment Algorithm – A Lesson on AI Bias

Abstract (Summary):In 2014, a team of machine learning specialists at Amazon began working on an innovative AI-powered recruitment tool. The goal of the project was to automate the resume screening process and identify top candidates for technical roles, including software engineering positions. The system rated applications on a scale from one to five stars—similar to how customers rate products on Amazon's platform. By 2015, the team discovered a serious issue: the algorithm was systematically biased against women applying for technical roles. The tool had been trained on résumés submitted to the company over the previous decade—most of which came from men, reflecting the male dominance in the tech industry. As a result, the algorithm "learned" that male candidates were preferable. The system penalized résumés that included the word "women's" (e.g., "women's chess club captain") and downgraded the scores of graduates from two all-women's colleges. It also favored applicants who used "masculine" verbs such as "executed" and "captured" in their résumés. Despite Amazon engineers' attempts to fix the bias, the company ultimately abandoned the project in 2018, acknowledging there was no guarantee the system wouldn't develop other forms of discrimination. This case became a crucial lesson for the tech industry on the risks associated with algorithmic bias and the importance of responsible AI development.

Theme of the Case Study:

- Artificial Intelligence in Recruitment
- Algorithmic Bias
- AI Ethics
- Gender Equality in Technology

Relevance to the Reader: Amazon's recruitment algorithm serves as a key case study in algorithmic bias in AI systems. It demonstrates how machine learning systems can not only reflect but also amplify existing societal inequalities when trained on biased historical data. At a time when AI is increasingly used in hiring processes (55% of HR professionals in the U.S. predict that AI will become a routine part of their work within five years), understanding the potential risks of algorithmic bias is vital for both employers and job seekers. The Amazon case shows that even the most technologically advanced organizations can unintentionally build discriminatory systems. The case study highlights the importance of diversity and inclusion in AI development. It illustrates how a lack of diversity in tech teams and training data can lead to systems that favor one demographic group at the expense of others. For professionals working in AI ethics, this case provides a concrete example of the challenges involved in ensuring algorithmic fairness. It underscores the need for rigorous testing of AI systems for potential biases before deployment, especially in high-stakes contexts like employment. This story is also relevant in the context of emerging AI-related regulations, such as New York City's anti-discrimination law on recruitment algorithms, which came into force in 2023. It emphasizes the need to establish standards and regulations to ensure that automated decision-making systems are fair, transparent, and accountable.

Case Study 3

AI-Driven Exam Integrity and Fairness: The YouTestMe Platform as a Case Study

Abstract (Summary):This case study examines YouTestMe, an AI-powered platform addressing challenges in online exam integrity and fairness.
Key innovations include:
• AI AutoMarking: Automated essay scoring using NLP to evaluate semantic coherence, structural integrity, and linguistic accuracy, reducing grader workload while maintaining consistency.
• Automated Proctoring: AI monitors test-takers for suspicious behaviors (e.g., window switching, face mismatch) with 98% detection accuracy, ensuring exam security.
• Fairness Perceptions: Studies show students perceive AI algorithms as fairer than human evaluators, particularly in formative assessments
• Transparency and post-evaluation explanations further enhance trust in AI- driven results. The platform aligns with Responsible AI principles by prioritizing data privacy, auditability, and human oversight. For example, proctoring footage is stored for optional human validation, blending AI efficiency with accountability.

Theme of the Case Study:
- Automated grading and proctoring in online exams
- AI fairness and bias mitigation in educational assessments
- Human-AI collaboration for transparent evaluations
- Responsible AI implementation in high-stakes testing

Relevance to the Reader:
Educators/Institutions: Demonstrates how AI can streamline assessment workflows while addressing biases. The study highlights that AI fairness perceptions improve when evaluation criteria are transparent and explanations are provided.
AI Developers: Offers a model for integrating ethical guardrails (e.g., bias checks in training data, user consent protocols) into EdTech systems, as seen in YouTestMe's GDPR-compliant data handling.
Policy Makers: Illustrates the importance of standards for AI proctoring and automated grading, emphasizing the need for human-in-the-loop oversight to mitigate risks like algorithmic hegemony.

This case study provides actionable insights for implementing AI in assessments responsibly, balancing efficiency with equity.

Case Study 4

National Youth Assembly on Artificial Intelligence – Youth Perspectives on AI Policy

Theme of the Case Study:

- AI Ethics & Regulation
- Youth Engagement in Policy
- AI Awareness & Literacy

Abstract (Summary):
In 2023, the National Youth Assembly on Artificial Intelligence convened young people aged 12 to 24 to discuss Ireland's National AI Strategy ("AI – Here for Good"). The initiative aimed to ensure that youth voices were included in shaping AI governance, policies, and education.

Key discussion areas included:
- Equality & Inclusion in AI – Concerns about bias and fairness in AI applications.
- Trustworthiness of AI – The importance of transparency and accountability in AI decision- making.
- AI Regulation & Public Awareness – The need for clear policies on AI use in schools, workplaces, and social media.

Participants proposed recommendations on ensuring AI supports youth empowerment, inclusion, and fair representation. Their perspectives were documented in a report used to inform government policy discussions.

Relevance to the Reader:
AI policies impact youth directly, but young people are often left out of decision-making.

This case study highlights the importance of youth engagement in shaping AI ethics and regulation.

Encourages educators and youth workers to promote AI literacy, ensuring that young people understand their rights and responsibilities in an AI-driven world.

Demonstrates how youth voices can contribute to AI policy—a practice that should be expanded across Europe.

Case Study 5

AI-Based Cyberbullying Interventions – Evaluating Youth Perspectives

Theme of the Case Study:

- Online Safety
- Bias in AI Moderation
- Youth Digital Wellbeing

Abstract (Summary):
This study investigated youth perceptions of AI-driven cyberbullying interventions on social media. Researchers at Dublin City University (DCU) tested AI-based proactive content moderation strategies designed to detect and limit harmful interactions online. The consultation involved young people aged 12 to 17, who evaluated these interventions through focus groups and online discussions.

The study found that while AI content moderation can reduce harmful interactions, youth participants expressed concerns over:
• False positives and censorship – AI incorrectly flagging harmless content as harmful.
• Lack of human context – AI struggling to understand nuances in humour and informal language.
• Privacy concerns – Uncertainty about who controls AI interventions and their impact on digital rights.

The research emphasised the need for youth participation in AI design, ensuring that content moderation aligns with young people's perspectives and digital realities.

Relevance to the Reader:

Many young people rely on AI-driven platforms for social interaction and education, making ethical AI implementation critical.

Cyberbullying is a growing concern, and while AI moderation can help, it must be transparent, fair, and accountable to users. This case study encourages youth workers and educators to discuss AI's role in digital safety, fostering critical thinking on algorithmic decision-making. • The study highlights the importance of engaging young people in the design and evaluation of AI interventions, ensuring that solutions are youth-centred and ethically sound.

Case Study 6

Responsible AI for Test Equity and Quality: The Duolingo English Test as a Case Study

Abstract (Summary):

This paper, forthcoming in the Handbook for Assessment in the Service of Learning, uses the Duolingo English Test (DET) as a case study to explore the critical role of Responsible AI (RAI) in high-stakes assessment. While AI offers efficiencies in item generation and scoring, it also presents risks, such as bias. The DET addresses these risks through four specific RAI standards—Validity and Reliability, Fairness, Privacy and Security, and Accountability and Transparency—demonstrating how these translate into practical implementation.

The paper details the development of these standards and their connection to broader RAI principles, providing concrete examples of how these practices ensure both test quality (valid score inferences) and equity (fairness for all test takers). The DET serves as a model for how companies can transparently and accountably integrate AI while mitigating ethical risks.

Theme of the Case Study:

- English proficiency test
- AI fairness and unbias in tests
- (Human in the loop) HiTL and AI
- Responsible AI

Relevance to the Reader:

This paper's relevance to the reader centres on the crucial role of Responsible AI (RAI) in ensuring both the quality and equity of AI-powered assessments. It uses the Duolingo English Test (DET) as a compelling case study, detailing how RAI principles are implemented in practice to mitigate the risks of AI while maximizing its benefits. This is valuable for anyone involved in assessment development, AI ethics, or the use of AI in high-stakes testing. The paper provides a detailed, practical example of how RAI principles can be implemented in a real-world, high-impact setting. This is crucial given the increasing use of AI in education and assessment. It also offers insights into the process of creating an AI-powered assessment, including item generation, scoring, and security measures. It highlights the importance of human-in-the-loop approaches for maintaining quality and equity. AI ethics and governance: The four RAI standards (Validity and Reliability, Fairness, Privacy and Security, Accountability and Transparency) presented serve as a model for organizations developing and deploying AI systems, particularly in high-stakes applications. The alignment with the NIST AI RMF further strengthens the framework's validity. The paper demonstrates how a systematic application of RAI practices can contribute to achieving both test equity (fairness for all test takers) and quality (valid score inferences). This is crucial for ensuring that AI-powered assessments are not biased or discriminatory.

Case Study 7

Gender bias in ads on prominent social media site.

Theme of the Case Study:

- Social Media
- Bias

Abstract (Summary):
Global Witness initiated some paid ads across Facebook in the following countries: UK, the Netherlands, France, India, Ireland, and South Africa. The ads focused on real job vacancies from various fields including pilots, hairdressers and psychologists. The results of who the ads were targeted at through the social media sites algorithm showed a clear gender bias relating to the job vacancy with 91% of individuals being shown the advert for the mechanic vacancy being male and 79% of individuals shown an advert for a pre-school teacher being female.

Upon further research in The Netherlands and France in collaboration with a feminist organisation, it was shown that 97% of individuals shown an advert for a receptionist vacancy were female.

As part of the mandatory selection process when creating an advert, Global Witness selected the objective of targeting those that Facebook thought would be most likely to click on the ad and it was only specified for each ad that it must be shown to adults who live in the country or had recently been in either country.

Relevance to the Reader:

This case study highlights the ingrained bias present in the algorithms that influence the content people see daily, particularly in job advertisements on social media. AI-driven systems, which determine ad placement, often reflect and reinforce existing societal biases, leading to disparities in how job opportunities are presented to different genders.

Social media has evolved beyond a platform for casual interaction; it now serves as a major source of information, support, and career guidance. However, algorithmic bias means that certain job roles may be disproportionately shown to men or women based on historical data rather than individual merit, limiting access to opportunities and reinforcing stereotypes.

Given the reliance of young people on social media for career-related information, it is crucial to raise awareness of this bias and equip them with critical thinking skills to recognise and counteract it. Encouraging transparency in AI systems and advocating for fairer algorithmic practices can help create a more equitable digital job market.

Sources

Case study 1:
•Axios: https://www.axios.com/2020/08/19/england-exams-algorithm-grading
•Wired: https://www.wired.com/story/skewed-grading-algorithms-fuel-backlash-beyond-classroom

Case study 2:
- Reuters – "Amazon scraps secret AI recruiting tool that showed bias against women" (2018): https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/
- LinkedIn – "Amazon's Biased Hiring Algorithm - A Wake-up Call for Tech" (2024) https://www.linkedin.com/pulse/amazons-biased-hiring-algorithm-2018-wake-up-call-tech-akshay-aryan-dfouc
- The Independent – "Amazon scraps 'sexist AI' recruitment tool" (2018) https://www.independent.co.uk/tech/amazon-ai-sexist-recruitment-tool-algorithm-a8579161.html
- Leoforce – "Lessons from Amazon's Sexist AI Recruiting Tool" (2024): https://leoforce.com/blog/amazon-ai-hiring-bias-case-study/
- Carnegie Mellon University – "Amazon Scraps Secret AI Recruiting Engine that Showed Biases Against Women" (2018) https://www.ml.cmu.edu/news/news-archive/2016-2020/2018/october/amazon-scraps-secret-artificial-intelligence-recruiting-engine-that-showed-biases-against-women.html
- ACLU – "Why Amazon's Automated Hiring Tool Discriminated Against Women" (2018): https://www.aclu.org/news/womens-rights/why-amazons-automated-hiring-tool-discriminated-against
- IMD – "Amazon's sexist hiring algorithm could still be better than a human" (2018): https://www.imd.org/research-knowledge/digital/articles/amazons-sexist-hiring-algorithm-could-still-be-better-than-a-human/
- Redress Compliance – "Amazon AI Hiring Tool: A Case Study in Algorithmic Bias" (2025): https://redresscompliance.com/amazon-ai-hiring-tool-a-case-study-in-algorithmic-bias/

Case study 3:

https://ceur-ws.org/Vol-3938/Paper_14.pdf

https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1221177/full

Punoševac, Mina & Nikolić, Anđela. (2024). Transforming Assessment Practices – A Case Study of YouTestMe. CEUR Workshop Proceedings

Case Study 4:

National Youth Assembly on AI Report (2023) Government of Ireland, Department of Enterprise, Trade and Employment https://enterprise.gov.ie/en/publications/publication-files/national-youth-assembly-on-ai-report.pdf

Sources

Case study 5:

- Milosevic et al. (2023). Effectiveness of Artificial Intelligence-Based Cyberbullying Interventions from a Youth Perspective
- Dublin City University (DCU) Research, Ireland

https://doras.dcu.ie/29411/1/milosevic-et-al-2023-effectiveness-of-artificial-intelligence-based-cyberbullying-interventions-from-youth-perspective.pdf

Case study 6:

https://arxiv.org/pdf/2409.07476
https://www.researchgate.net/publication/383985608_Responsible_AI_for_Test_Equity_and_Quality_The_Duolingo_English_Test_as_a_Case_Study

- Burstein, Jill & LaFlair, Geoffrey & Yancey, Kevin & von Davier, Alina & Dotan, Ravit. (2024). Responsible AI for Test Equity and Quality: The Duolingo English Test as a Case Study.

Case study 7:

https://globalwitness.org/en/campaigns/digital-threats/new-evidence-of-facebooks-sexist-algorithm/

- Published on the 12 June 2023
- Updated on the 13 February 2025

Conclusion

As we have explored throughout this guidebook, the integration of artificial intelligence throughout the different areas of society; education, employment, media and more, offers both transformative potential and meaningful challenges. The case studies presented demonstrate the diverse and creative ways AI is being harnessed.

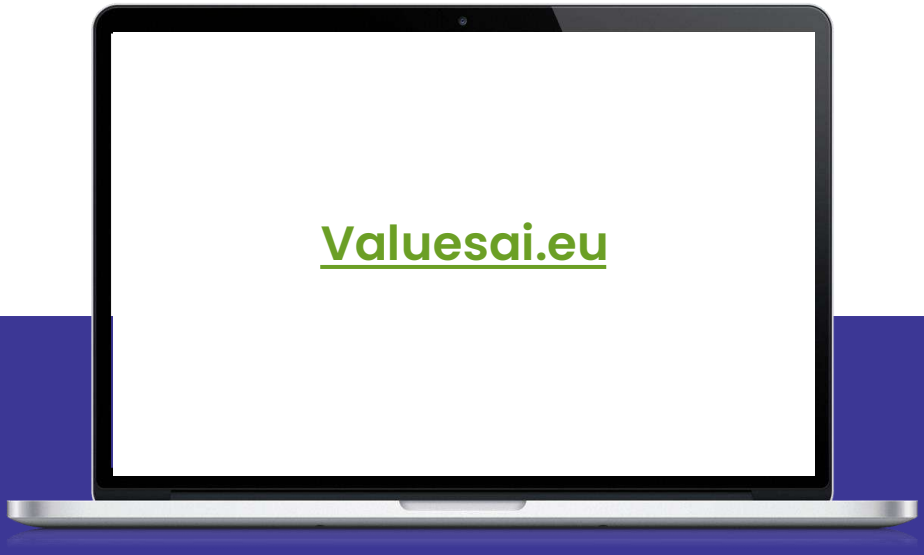
What unites these examples is a common thread of innovation driven by the needs and voices of young people. They show that when AI is used thoughtfully and ethically, it can empower young people to take more ownership of their education, employment and personal lives, whilst fostering critical thinking about the technology itself.

However, these advances also underscore the importance of digital literacy, responsible use, and equity of access. Educators, youth workers, and policy-makers all have a role to play in guiding the use of AI in ways that prioritise human connection, ethical practice, and lifelong learning.

This guidebook is not an endpoint but an invitation—to continue exploring, questioning, and co-creating a future where young people are not just users of AI, but active shapers of its role in education and society.

AI Skills, Powered by Values

values



www.valuesai.eu

follow our journey

Co-funded by the European Union